



UNITED NATIONS  
UNIVERSITY  
**UNU-WIDER**

WIDER Working Paper 2020/151

## **By choice or by force?**

Uncovering the nature of informal employment in urban Mexico

Robert Duval Hernández\*

November 2020

**Abstract:** Using a special module of the 2015 Mexican Labour Force Survey with information on workers' preferences for jobs with social security coverage, I estimate that 80 per cent of informal workers in large urban areas would prefer to work in a job that provides them with such coverage. A discrete choice econometric model which distinguishes between wanting a formal job and the probability of getting one shows that schooling increases the chances of being hired in formal employment and of having higher earnings in it. Women with greater responsibilities at home are less likely to want formal employment, and they also face a lower probability of being hired in such jobs. The findings indicate the segmentation of Mexican labour markets and the rationing of formal jobs, together with the existence of workers who voluntarily participate in informal employment. However, the estimated fraction of involuntary informal workers is quite high.

**Key words:** informal employment, labour markets, segmentation, rationing

**JEL classification:** O17

**Table and figures:** at end of the paper

**Acknowledgements:** This paper is part of the project 'Transforming informal work and livelihoods' of the United Nations University World Institute for Development Economics Research (UNU-WIDER). I would like to thank Gary S. Fields and two anonymous referees for their valuable comments. The opinions and views here expressed are my own and do not represent those of my employer or of UNU-WIDER.

**Note:** the online Appendix is available [here](https://www.wider.unu.edu/publication/choice-or-force) (<https://www.wider.unu.edu/publication/choice-or-force>).

---

\*Faculty of Economics and Management, Open University of Cyprus, CIDE, Mexico and Institute of Labor Economics (IZA), Bonn, Germany; [robert.duval@ouc.ac.cy](mailto:robert.duval@ouc.ac.cy)

This study has been prepared within the UNU-WIDER project [Transforming informal work and livelihoods](#).

Copyright © UNU-WIDER 2020

Information and requests: [publications@wider.unu.edu](mailto:publications@wider.unu.edu)

ISSN 1798-7237 ISBN 978-92-9256-908-2

<https://doi.org/10.35188/UNU-WIDER/2020/908-2>

Typescript prepared by Lesley Ellen.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

## 1 Introduction

A central question in the study of labour markets in the developing world is whether its large share of informal employment is mainly composed of workers who cannot find a better employment option in the formal economy, or whether this sector is formed by workers who voluntarily seek this type of employment given the incentives in the market.

Broadly speaking, traditional theories of dualistic labour markets conceive informal employment as a second-best option for workers who cannot find a formal job, and who cannot afford to remain unemployed while searching for a good employment option (see, for instance, the seminal model of Fields 1975). In contrast, a different strand of the literature has argued that workers optimally choose informal employment because, given their individual characteristics, they obtain a higher utility/payoff in that type of work (see, for instance, Maloney 1999). This can arise when workers prefer to be informally employed because they will pay lower taxes, face less regulations, or have more flexible contract arrangements than if they were formally employed.

Despite the above opposing views, the literature recognizes that, in practice, informal employment is heterogeneous. It is composed of workers who participate in it voluntarily and those who are there because they cannot find a formal job (Fields 1990; Maloney 2004). While there is agreement about the heterogeneous nature of informal employment, little is known about how many informal workers are voluntarily in informal employment and how many simply cannot find a better job elsewhere. Little is also known about the characteristics of these different types of workers.

This gap in our knowledge arises because, in most of the available survey data, workers are not asked about their preferred type of employment and, for this reason, researchers have devoted a substantial amount of energy in testing for the existence of rationing in labour markets through: i) the estimation of complex structural econometric models of sector allocation (e.g. Günther and Launov 2012; Magnac 1991); ii) the analysis of sector transitions over the business cycle (e.g. Bosch and Maloney 2010; Fiess et al. 2010); and iii) the use of ancillary information about the characteristics of informal employment in order to identify those workers who perform successfully in it and, thus, are more likely to participate in it voluntarily (e.g. Gindling and Newhouse 2014).

The main contribution of this paper is to use a special supplement of the 2015 Mexican Labour Force Survey (INEGI 2015a) which directly asks urban workers about their valuation of jobs with social security coverage. Since, in Mexico, having social security coverage in a job is the defining characteristic of formal wage employment, this piece of information gives us a proxy measure of workers' preferences for these types of jobs.

This variable is used to address important questions in the informality literature, such as: What fraction of informal workers would rather be formal employees? What individual characteristics increase the probability of applying to and of being hired in a formal job? What are the earnings that different types of informal workers would obtain if they had formal employment?

Our analysis indicates that around 80 per cent of the respondents who lack social security coverage would prefer to have a job with such benefits, even if that entailed having to pay the corresponding contributions for them. Although one cannot determine whether this reflects a preference for social security benefits alone or, more generally, a preference for the entire set of characteristics that accompany a formal job, the figure indicates that a large fraction of the urban informal workers

in Mexico are not voluntarily so. Even among those informal workers with better jobs, i.e. among the ‘upper-tier’ informal workers, about three-quarters of them would prefer to have a job with coverage.

A discrete sector choice model that distinguishes the worker’s decision to apply for a formal job from the formal sector employer’s hiring decision is estimated. The results of this model reveal that women living in households with a higher fraction of dependents are less likely to want formal employment, likely because a heavier burden of household duties requires jobs with flexible schedules, such as the informal ones. In fact, this negative effect of dependents on the probability of applying for formal jobs disappears once other adult females (who can help with care activities at home) are present in the household. For males, the opposite effect is observed. Namely, a higher fraction of dependents is positively associated with the probability of males applying to formal jobs. This finding is consistent with a more traditional role of men as the main breadwinners in the household.

The number of years of schooling significantly increases the probability of being hired in a formal job, but workers still enrolled in school face a substantial hiring penalty. Finally, marital status affects the probability of being hired differently for men and women. Married men are more likely to be hired in formal employment compared to single males, while the exact opposite occurs for married females. Again, given the unequal division of labour at home in Mexico, this reflects that formal sector employers prefer workers who have fewer household responsibilities.

Selectivity-adjusted earnings equations are estimated for three different types of workers: formal, voluntary informal, and other informal workers. The estimations show a large gender wage premium, and married women who voluntarily work in an informal job face large earnings penalties compared to single females. The returns to education for formal workers are around 15 per cent, and between 4 per cent and 8 per cent for the voluntary informal. The earnings of involuntary informal workers show no statistically significant relation to education.

Overall, the empirical findings confirm the view that informal employment is formed of a heterogeneous group of workers, some of whom participate voluntarily in it while others do so because of a lack of better options. However, contrary to previous suggestions in the literature, the fraction of involuntary informal workers is quite high.

Broadly speaking, the findings of this paper highlight two main factors that limit the number of workers employed in formal jobs. The first is related to household demographics and the division of housework. In particular, women who have a higher burden of work at home are less likely to seek formal employment and are less likely to be hired in formal jobs. The second is related to human capital, as higher levels of schooling increase the chances of being hired in formal jobs and of obtaining higher earnings in them.

While the econometric estimations cannot be used to predict the consequences of policy changes that occur at the aggregate level, they nevertheless indicate some of the dimensions that need to be considered when designing reforms that seek to encourage the growth of better-paying formal employment. Also, this study highlights the usefulness of incorporating direct information about the valuation of different types of jobs in the study of labour markets in developing countries.

Before delving into the analysis of the survey supplement which leads to the key results in this paper, Section 2 presents a broad overview of the structure of urban labour markets in Mexico based on data from the Labour Force Survey. In Section 3, the data on the valuation of jobs with social security benefits is presented together with its main descriptive statistics. Section 4 discusses

the econometric methodology for the *ceteris paribus* analysis of job allocation and earnings, while the results of these estimations are presented in Section 5. Section 6 concludes.

## 2 A descriptive view of Mexican urban labour markets

This section presents an overview of the structure of Mexican urban labour markets, with an emphasis on the heterogeneity of the different segments it comprises. More precisely, it presents key descriptive statistics for six types of workers: formal wage employees, formal self-employed, upper-tier wage employees, upper-tier self-employed, lower-tier wage employees, and lower-tier self-employed. As emphasized in Alaniz et al. (2020) and Fields (forthcoming), this taxonomy has three defining dimensions, namely: i) it distinguishes between formal and informal status; ii) it distinguishes between wage employees and the self-employed; and iii) it creates two tiers (upper and lower) within the informal workers category, where the upper tier is more likely composed of informal work with higher remunerations with barriers to entry and, therefore, this tier is more likely to be a segment of the market where workers participate voluntarily.

The operationalization of these categories for the Mexican case is summarized in Table 1. While there is no universal definition of what constitutes an ‘informal job’, in the case of Mexico two dimensions have been used as defining characteristics to establish whether a job is formal or not.<sup>1</sup> In the case of wage employees, the defining criterion for distinguishing between formal and informal workers is whether the job offers social security coverage. For the self-employed, the formality status of a business is determined by whether it is registered with tax authorities and has fixed work premises (see, for instance, INEGI 2014).<sup>2</sup> This second criterion is relevant because Mexican law does not mandate self-employed individuals to register themselves for social security coverage.<sup>3</sup>

The self-employed in the upper tier are those who: i) are employers (i.e. with employees of their own); or ii) have a place of work with fixed premises; or iii) voluntarily enroll in a social security scheme; or iv) work in an occupation that requires post-secondary education.<sup>4</sup> Similarly, wage employees in the upper tier are those in skilled occupations, those with a permanent contract, or those who, although they do not receive social security coverage, receive other fringe benefits at

---

<sup>1</sup> See, for instance Levy (2018). See also Fields (2011) and Rupert Bulmer (2018) for thoughtful discussions of the many definitions that the term ‘informality’ takes in the literature.

<sup>2</sup> In addition, all employers in agriculture with at least one employee are considered formal. In this paper, this type of employer is less relevant as it focuses its analysis only on large urban areas. The criteria for distinguishing formal work are those used by the Mexican Statistical Agency, INEGI (Instituto Nacional de Estadística y Geografía 2014) in agreement with the guidelines set by the International Labour Organization (ILO). It should be noted that under the classification adopted by INEGI, there can be informal workers employed in the formal sector, but not the other way around. A business is considered part of the formal sector if it is a registered entity for legal and tax purposes. For more details, see INEGI (2014).

<sup>3</sup> They can, however, register voluntarily in a government-sponsored social security scheme. However, the vast majority of the self-employed do not do so.

<sup>4</sup> I use the Labour Force Survey to find the 4-digit occupations where the majority of workers have post-secondary or vocational education. These include most of the managerial, professional, and technical occupations of the ISCO classification, and a few select others.

work. All other informal workers are classified in their corresponding lower tier, depending on whether they are wage employees or self-employed.<sup>5</sup>

Table 2 presents descriptive statistics for a representative sample of the different worker categories outlined above, as well as for the unemployed and those out of the labour force.<sup>6</sup> The table shows that about half of the employed population are formally employed, almost a third are in lower-tier informal jobs, and the rest are in upper-tier informal jobs. Most formal workers are wage employees, as the self-employed represent only one-ninth of the formal employed population. In contrast, one in three informal jobs are in self-employment, irrespective of tier.

The self-employed are, on average, ten years older than wage employees, and males overwhelmingly participate in formal employment and upper-tier self-employment, while the gender distribution is almost equal in the other segments of the market and among the unemployed. Schooling decreases monotonically as one moves down the job ladder from formal employment to upper-tier and, lastly, to lower-tier informal employment. However, the unemployed are not the least-educated workers, and they have comparable education levels to those of workers in the upper-tier informal segment. This shows that, in Mexico, the unemployed are not the worst-off workers but, rather, a group that can ‘afford’ to keep searching for better jobs without having to enter into informal employment as a last-resort option. Finally, very few self-employed are still enrolled in school, while this share is 10 per cent among the informal wage employees.

In addition to the above statistics, the table also presents selected statistics on demographic variables at the household level because, as will be seen later, these are relevant for workers’ decisions to seek formal jobs. Specifically, the statistics show that informal wage employees are more likely to be single than the rest of the employed population. On average, the dependency ratio at the household level, i.e. the fraction of household members younger than 18 years or older than 65 years relative to the number of members between the ages of 18 and 65, ranges between 0.52 and 0.6, except for lower-tier informal workers and individuals out of the labour force, for whom this ratio is higher. This variable is relevant because having a larger number of dependents at home will require more hours devoted to their care and a greater income to sustain them.<sup>7</sup>

Table 2 also includes statistics on job characteristics for each segment of employment. While these variables are not directly used in the analysis in the coming sections, they serve to illustrate the heterogeneity between these segments of the labour market.

In terms of occupation and industry, certain categories such as managerial, professional, and clerical occupations, as well as jobs in manufacturing, business services, education, and health are over-represented in the formal and upper-tier segments of the market. In contrast, jobs in personal services are over-represented in informal employment, especially at the lower tier. As expected,

---

<sup>5</sup> All unpaid employees are classified as lower-tier informal wage employees.

<sup>6</sup> The data used for this table is the Labour Force Survey (Encuesta Nacional de Ocupación y Empleo, ENOE) (INEGI 2015b) and I use geographical and temporal criteria to make it comparable with the module that is analysed in the remaining sections of the paper. Large urban centres are those with a population of 100,000 or more. Comparable statistics for the entire country are available upon request from the author. To the extent that informality has been traditionally conceived as being in urban settings, little is lost through this narrower geographical focus. Furthermore, the Module which enquires about preferences for social security in the next section is representative only at the level of large urban centres. All the estimates in this paper are weighted using sampling weights.

<sup>7</sup> The number of adult females in the household is homogeneous across groups.

most public sector workers are formal, but even this sector hires some wage employees informally. Finally, as previously mentioned, only a small fraction of the formal self-employed voluntarily choose to contribute to a social security scheme.

The last row of Table 2 presents the average earnings for each group of workers. These numbers are reproduced in Figure 1. Two patterns are visible in the figure. First, earnings fall as one moves down the job ladder and, second, within each segment (formal, upper-tier, and lower-tier) the self-employed have higher earnings than the corresponding wage employees in their segment. There may be two reasons for this last pattern: first, on average, the self-employed are older and thus have more experience; and, second, their reported earnings likely capture payments to both labour and capital productive factors, especially for those self-employed operating in the formal sector.

To close this descriptive section, Table 3 presents panel data evidence on the segment transitions of workers over a period of one year for those who were employed during both the initial and final periods.<sup>8</sup> Several interesting findings arise from this table. First, formal wage employees have the highest share of stayers among all groups (87.4 per cent), reflecting the greater job security enjoyed by these workers and the fact that they are at the top of the job ladder and are therefore less likely to change jobs willingly. In contrast, the greatest turnover is found among upper-tier informal workers, with about 34 per cent staying in this segment after one year. Of the upper-tier movers, about half move down the job ladder to lower-tier informal employment. The fact that lower-tier informal workers have a large share of stayers (around 58 per cent), second only to formal wage employees, implies that not only do these workers have the worst jobs but they also have a small chance of moving out to a better segment of the market.

While the analysis of transitions provides useful information about how likely it is for some workers to be trapped in a vicious cycle of poor-quality employment, it is not conclusive on the question of how many of the stayers do so voluntarily rather than because of a lack of opportunity elsewhere. To answer this, it is necessary to have information about workers' preferences for certain types of jobs. In the remainder of the paper, I analyse a special module of the 2015 Labour Force Survey (INEGI 2015a) which contains this type of information.

### **3 Data on valuation of social security coverage**

In the second quarter of 2015, the Mexican Labour Force Survey was supplemented by a module inquiring about the employment trajectories of workers in large urban areas and their contribution to and valuation of social security protection.<sup>9</sup> The MOTRAL module (after its acronym in Spanish) was applied to a representative sample of workers aged 18 to 54, living in large urban centres who were either employed or had previous labour market experience. This target population represented around 90 per cent of the labour force in large urban centres and 60 per cent of the overall urban labour force in 2015. The individuals interviewed in the module also

---

<sup>8</sup> The full transition matrix, including transitions into unemployment and out of the labour force, is available from the author upon request.

<sup>9</sup> The supplementary module is the *Módulo de Trayectorias Laborales, 2015 (MOTRAL)* and its data can be publicly accessed online (see INEGI 2015a). A similar module was also applied in 2012, but in that edition the key variable used in the analysis was not included.

answered the questions in the Labour Force Survey, and the two datasets can therefore be linked, as is done here, in order to have a richer set of variables.

This module included the following key question: Do you think it is better to have a job with social security, even if you have to make payments to be eligible for it?<sup>10</sup> This question is central to the analysis because most labour surveys contain information on the sector of employment, but they do not collect information on the types of jobs workers value. Without this piece of information researchers have no choice but to try to infer through indirect methods what fraction of the informal workforce is so because of a lack of options rather than by choice.

Given that having social security coverage is the defining characteristic of formal wage employment in Mexico, this variable can be used as a proxy for the value workers give to this type of work. Linking this variable to information on the *actual* type of job (formal or informal) can help us to approximate the fraction of involuntary informal workers.

### 3.1 Descriptive statistics

Table 4 shows the percentage of the population who report preferring a job covered by social security protection and the main reason for valuing this protection among those who prefer jobs with social security coverage. The first column shows that across all groups the vast majority respond affirmatively to the question of whether they prefer to have a job with coverage.

The group which displays the highest proportion of positive responses is the unemployed, with rates of almost 93 per cent. Of the employed population, the most positive answers are found among the lower-tier self-employed and the formal wage employees, with affirmative answers from about 85 per cent of the respondents. In contrast, the lowest share of workers who answer this question affirmatively is found in the upper-tier informal, where ‘only’ about 74 per cent of workers answer affirmatively. The table also shows that among those respondents who prefer jobs with coverage, the most-valued attributes of social security protection are usually health and pension benefits.<sup>11</sup>

Table 5 compares key sociodemographic characteristics of workers depending on whether they have social security coverage in their job and on whether they would prefer to have a job with coverage.<sup>12</sup> The table shows that the age and gender composition is more or less homogeneous across groups except for workers with coverage who do not value their social security benefits (in column 2), who are predominantly male. Workers with coverage are more educated than those without coverage, and within each market segment (with coverage or without coverage) respondents who do not value social security have higher levels of schooling than those who value it. In addition, workers without coverage are more likely to still be enrolled in school. Workers without coverage who do not value social security benefits are less likely to be married, have a higher number of dependents at home, and have the lowest earnings of all groups. In contrast, the

---

<sup>10</sup> The original question reads: ‘¿Considera que es mejor tener un empleo con seguridad social, aunque tenga que realizar pagos para tener derecho a ella?’.

<sup>11</sup> In Mexico, formal workers have very limited unemployment benefits and, probably because of this, the access to such benefits was not listed in the MOTRAL questionnaire as a reason for valuing social security coverage.

<sup>12</sup> Table A1 in the Appendix presents a more comprehensive set of descriptive statistics of the MOTRAL sample used in the estimation of econometric models.



group that exhibits higher average earnings is the workers with coverage who do not value social security coverage.

Taking advantage of the fact that the Labour Force Survey is a short-lived rotating panel, and that the MOTRAL module can be linked to it, one can also explore, even if in a limited way, the employment transitions of workers with different preferences for jobs with coverage. In particular, the individuals in the MOTRAL module participated in the 1<sup>st</sup> Quarter interview of the 2015 Labour Force Survey but were interviewed for the module during the 2<sup>nd</sup> Quarter of 2015. Assuming that their valuation of jobs with coverage remained unaltered between these two quarters, one can then compare the one-quarter transition matrices across eight categories (the aforementioned six employment categories, plus the unemployment and out of the labour force categories) separately for workers according to their preferences for jobs with coverage.<sup>13</sup> The results of these transitions are presented in Table 6.

Comparison of the highlighted column in panels A and B of the table shows that in all but two employment categories the transitions into formal wage employment are higher among the individuals who prefer a job with coverage. This likely reflects a higher search intensity for jobs with coverage among individuals who value such jobs. Although these stratified transition matrices are limited by the proximity of the periods considered and by the small sample size of the MOTRAL module, they serve to illustrate the utility of enriching a traditional transition analysis with information about the valuation of certain types of jobs.

### 3.2 Voluntary versus involuntary informal employment

As previously mentioned, this paper's key contribution is to enrich the study of informal labour markets in Mexico by analysing not only the traditional variables on employment and wages by sector but also the valuation workers attach to jobs with social security coverage.

While the question of whether it is better to have a job with social security benefits is a new valuable piece of information, its interpretation requires careful thought. In particular, there are two ways to interpret the answers. The first, which I call the *ceteris paribus* answer, compares jobs that are similar along all relevant dimensions (e.g. wages, hours, job security, etc.) except for the presence or not of social security benefits. An alternative way of answering is to compare a 'typical' formal wage job with the one the respondent currently has, and to report which is considered better. I call these *unconditional* answers.

In the case of formal wage employees, I assume they prefer formal jobs to informal ones, irrespective of their valuation of social security benefits. This means that the 16 per cent who do not value their social security benefits (see Table 4), still prefer formal to informal employment, as otherwise they would voluntarily move to an informal job. In other words, they remain in formal employment because of other job qualities such as higher wages, greater job security, and generally better working conditions.<sup>14</sup> This interpretation is corroborated by the descriptive evidence in the

---

<sup>13</sup> As the Labour Force Survey interviews individuals over five consecutive quarters, one could also construct transition matrices over longer periods. However, due to i) the rotating structure of the panel, ii) the relatively small sample size of the MOTRAL module, and iii) sample attrition, the sample size on matrices over longer horizons makes the analysis unreliable.

<sup>14</sup> These workers answer the question about the valuation of jobs with coverage in a *ceteris paribus* way. That is, they prefer a job without coverage, but they remain formally employed because of other characteristics that make it attractive to remain formal. This logic applies even if one considers that workers might be choosing their sector of employment with dynamic considerations in mind. For instance, workers might choose to be formally employed in

previous section, where column 2 of Table 5 shows that workers with coverage who do not value social security coverage are, on average, more educated than the others and have the highest average earnings of all groups.

In the case of informal workers, I assume that those who consider it better to have a job with social security coverage are involuntarily employed in their current job. That is, I interpret their answers as expressing an *unconditional* comparison between jobs with coverage and their current employment. Based on the figures in Table 5, it can be seen that under this interpretation, involuntary informal workers make up 38 per cent of the employed, and the voluntary informal comprise 9 per cent of the employed.

On average, formal wage jobs are better than informal wage jobs across several dimensions, such as pay, job security, and work conditions. Therefore, even if an informal wage worker were to interpret the question in a *ceteris paribus* way, more often than not, one would still expect such a worker to have an overall preference for a formal wage job.<sup>15</sup>

The risk of misclassifying the informal self-employed as involuntary is greater if they answer the question in a *ceteris paribus* manner. In this case, they might state that they would prefer a job similar to their current one, but with social security benefits, but they might be unwilling to give up being independent workers in order to become wage employees. The fact that very few self-employed voluntarily enrol for social security coverage in a government-sponsored programme indicates that few of them value social security *given* their current employment characteristics. Yet, the responses recorded in Table 4 suggest that they would prefer a job with coverage *if* that job had characteristics similar to the ones typically found in formal sector jobs. In other words, these workers are answering the MOTRAL questionnaire doing an *unconditional* comparison between jobs. To explore this issue in more detail, I use ancillary information from a survey on micro-entrepreneurs conducted in 2012.

The last edition of the Encuesta Nacional de Micro-Negocios (ENAMIN, or National Survey on Micro-enterprises) was conducted in 2012 (INEGI 2012). This nationally representative survey collected detailed information on the characteristics and operations of micro-enterprises in non-agricultural sectors.<sup>16</sup> The survey asked employers and the self-employed whether they would accept a wage job with a similar income to their current one but in which they had access to social security and/or pension benefits. In addition, entrepreneurs were asked their reservation monthly earnings for them to close their business and become wage workers.<sup>17</sup> As the wording of these questions is more precise than the one in the MOTRAL module, they can be used to descriptively explore the preferences of such workers, even if the sample does not represent the majority of the

---

order to gain a better income and save to become entrepreneurs later on in their career (see Maloney 2004 for some evidence on this type of life-cycle pattern).

<sup>15</sup> It is, however, conceivable that some informal wage workers prefer to remain informal because formal wage jobs are usually less flexible in terms of work schedules. Therefore, if informal workers who prefer flexible schedules were to answer the question in a *ceteris paribus* way, one might misclassify them as involuntary informal.

<sup>16</sup> A firm is considered to be a micro-enterprise if it employs ten workers or fewer, or if it is in manufacturing and employs fewer than 16 employees.

<sup>17</sup> The specific questions are: '¿Aceptaría un trabajo asalariado con un ingreso similar a lo que actualmente gana en su negocio o actividad, pero donde usted tuviera acceso a los servicios del IMSS?', '¿Cambiaría su negocio o actividad actual, si le ofrecieran un trabajo con un ingreso similar al que tiene actualmente y además tuviera derecho a una pensión?' and '¿Cuál es el salario mensual por el que usted aceptaría dejar su negocio o actividad?'

informal workers (as it excludes some informal entrepreneurs with large firms and all informal wage employees).

Table 7 presents the responses to these questions.<sup>18</sup> The first column shows that around 60 per cent of the respondents would accept wage employment with social security coverage if it paid the same income as the one they currently receive, and this proportion increases monotonically as we move down the job ladder from the formal self-employed to the lower-tier informal. The second column shows the proportion of respondents who would not accept becoming wage employees. This ranges between 32 per cent and 43 per cent and, as expected, increases as we move up the job ladder. Finally, the last three columns report the remaining proportion who would not accept wage employment with social security if it paid their current income, but who would accept it for a different level of pay. In particular, between 3 per cent and 10 per cent would close their micro-enterprise and become a wage employee for a different salary, but the median level of wage required for them to make this transition is several multiples of the median earnings paid in formal wage jobs.

As previously mentioned, the above numbers are not strictly comparable with those reported in Table 4 for the MOTRAL sample, as they come from different years and the ENAMIN sample does not include all self-employed. However, they provide a useful benchmark to compare with the responses reported in the main survey used, MOTRAL. In particular, both tables reveal that a large proportion of the self-employed would prefer to have formal wage employment. However, the proportion of self-employed in the 2015 MOTRAL module who consider formal wage jobs to be better is higher than in the 2012 ENAMIN. While the differences in years and sample coverage might account for part of this difference, it is likely that the different wording of the questions affected the answers.<sup>19</sup> Keeping this caveat in mind, I proceed using the MOTRAL sample for the rest of the analysis because of its broader coverage.

At this point, it is important to note that, following the classification adopted for who is a voluntary informal worker and who is not, there is no immediate one-to-one association between being voluntarily in informal employment and being in the upper tier previously defined. As previously shown in Table 4, a large proportion of the upper-tier informal workers still consider jobs with coverage to be preferable. Furthermore, Table 5 showed that voluntary informal workers have the lowest earnings of all groups, indicating they are not the ‘most successful’ informal workers. As will be seen in the coming sections, the valuation of jobs with coverage is strongly influenced by the division of labour at home, and not necessarily by success in informal employment.

While the idea that upper-tier workers are more likely to be voluntarily in informal employment has intuitive appeal, there is no logical necessity for this correspondence to occur. Some workers may have a comparative advantage in being employed formally, but they may also have an absolute advantage in both types of employment. Therefore, if they are unable to obtain a formal job, they are likely to find an upper-tier informal job. In addition, being a voluntary informal worker only implies that these workers are better off in informal employment relative to how well off they

---

<sup>18</sup> The table was estimated using a sample comparable to the MOTRAL one, namely workers at large urban centres, aged between 18 and 54, and with previous labour market experience.

<sup>19</sup> It is conceivable that, while some self-employed workers stated that formal wage jobs were ‘better’, when asked whether they would be willing to shut down their current business, they replied in the negative.

would be in formal employment. It does not mean that they are better off than other informal workers.

Going back to the question of the proportion of workers that can be deemed ‘involuntary informal’, it is clear that this proportion is quite high irrespective of the employment group. In particular, based on the figures reported at the bottom of Table 5, and given the classification used to determine who is a voluntary informal worker and who is not, around 80 per cent of informal workers are involuntarily employed.<sup>20</sup> This high proportion of involuntary informal workers contrasts with the view put forward by a strand of the informality literature which considers Mexican labour markets to be mainly composed of voluntary workers (see, for instance, Bosch and Maloney 2010; Maloney 1999, 2004).

This literature reaches this conclusion mainly based on an analysis of the patterns of sector transitions over the business cycle rather than from the direct measurement of workers’ stated preferences, as this paper does. Sector transitions over the business cycle, while interesting on their own, only provide indirect evidence about the preferred sector of employment of a given worker. Furthermore, sector transitions, by definition, do not tell us anything about the preferred jobs of stayers, as they may remain in their employment either because they are satisfied with it or because they face significant barriers to moving to a better job elsewhere.<sup>21</sup>

Our results also contrast with the argument that part of the reason informality is so high in Mexico is because many workers find the benefits associated with social security not worth the taxes that have to be paid to obtain them (Levy 2008). While it is possible that some informal workers do not value such benefits *given their current employment*, the answers in the MOTRAL module and the ENAMIN survey indicate that a large proportion of informal workers would prefer to be employed as wage employees in a formal job with better pay and working conditions.

In addition to the above studies, other papers have tried to estimate the proportion of involuntary informal workers in Mexico using structural econometric methods. Two examples are the papers by Duval-Hernandez and Smith (2010) and Alcaraz et al. (2015). Both use variants of discrete choice models that allow for rationing of formal jobs under a context of partial observability because they lack information on the preferred sector of workers.<sup>22</sup> The estimates of the

---

<sup>20</sup> This estimate is obtained from the figures reported in columns 3 and 4 of Table 5. There, it is reported that 38.3 per cent of the sample are involuntary informal, and 9.3 per cent are voluntary informal. Therefore, of the total informal population (which represent 47.6 per cent of the sample), 80 per cent are involuntary informal (i.e. 38.3/47.6). The same proportion of involuntary informal workers would arise if we defined formality not just by social security coverage status, but if we also included the formal self-employed among them. See Table A1 in the Appendix for details.

<sup>21</sup> Another potential reason for the divergence in results may be that the sample focuses on a segment of the employed population with a strong attachment to the labour force (they are aged between 18 and 54, and are either employed or had previous labour market experience). Hence, it is possible that if one were to extend the MOTRAL survey among labour market participants without previous work experience, or outside the 18–54 age range, then the share of involuntary informal workers might fall. However, this seems implausible as a comparison of the descriptive statistics between the entire Labour Force Survey and the MOTRAL module shows both samples to be very similar, except that the MOTRAL module has a lower share of males, is slightly younger, and has a larger share of formal workers. This comparison is available upon request from the author.

<sup>22</sup> Other papers have tried to test segmentation by comparing formal and informal wage equations after correcting for self-selection into each sector. However, the methods used to correct for sample selectivity are often based on sector choice models that assume free entry into formal employment. This is problematic as free choice among sectors is precisely the issue these papers try to test (see, for instance, Marcouiller et al. 1997 for the case of Mexico).

proportion of involuntary informal workers vary widely between these studies, and sometimes even within a given paper, depending on the sample analysed. For instance, using a sample of males between the ages of 20 and 60 in large metropolitan areas in 2005, Duval-Hernandez and Smith (2010) estimated that between 27 and 60 per cent of informal workers were involuntary workers. In contrast, Alcaraz et al. (2015) estimated that proportion to be between 10 and 20 per cent using a related model in a sample of males in localities with a population of 15,000+ in 2014.<sup>23</sup> The fact that these estimates vary widely depending on the econometric specification and the sample under analysis highlights the potential danger of trying to infer the proportion of involuntary informal workers using indirect methods.

To conclude this section, it is worth mentioning that a few studies in other countries have exploited questions similar to the one used in this paper, where workers are explicitly asked whether they would prefer to have a formal job. Using a special module of a household survey in Brazil in 1990, Soares (2004) estimated that around 70 per cent of informal wage workers and 30 per cent of the self-employed would rather be employed as formal wage workers. Also, using data for Chile in 2009, Contreras et al. (2017) estimated that around 35 per cent of the self-employed would rather work as wage employees. The proportion of involuntary informal wage workers in Brazil is similar to the one found in the MOTRAL survey in this paper. However, the proportions of involuntary self-employed in Brazil and Chile are substantially lower than those found in Mexico either in 2015 in the MOTRAL module or in 2012 in the ENAMIN survey.

In general, there is room for more extensive use of these types of counterfactual questions to complement the information regularly collected by the labour surveys. For example, Ulysea (2011) argues that the analysis of traditional variables such as wages, job transitions, and durations is insufficient to identify whether labour markets are segmented or integrated. Therefore, incorporating information about preferences seems a fruitful way to enhance our understanding of the workings of the labour markets in developing countries.

The next section presents the econometric models used in this paper to exploit the extra information provided by the MOTRAL module, with the goal of identifying the factors associated with applying for and being hired in formal employment.

#### **4 Econometric methodology**

As previously emphasized, this paper's main contribution is its analysis of a dataset that contains information about the preferred type of job for a representative sample of workers. By having information that distinguishes between the desired and the actual types of job, one can estimate econometric models of sector assignment that disentangle applications from hiring decisions for formal jobs.

To set the notation,  $V_i^a$  denotes the utility of worker  $i$  of applying and being employed in a formal job, and  $V_i^h$  denotes the corresponding propensity of a formal sector employer of hiring this worker  $i$ . Let us assume that such propensities depend on observable characteristics of the

---

<sup>23</sup> Both papers limit their analysis to males with positive earnings to focus on a group of workers with a closer attachment to the labour force, and to avoid having to model labour force participation decisions for females.

worker,  $(Z_i, X_i)$ , as well as on a set of unobservables,  $(u_{ai}, u_{hi})$ .<sup>24</sup> In particular, we assume that these components are related by the following system of equations:

$$V^a = Z\gamma_a + u_a \quad (1.1)$$

$$V^h = X\gamma_h + u_h \quad (1.2)$$

where  $(\gamma_a, \gamma_h)$  is a pair of vectors of unknown parameters.<sup>25</sup> To estimate the parameters from this model, we assume the vector of unobservables  $(u_a, u_h)$  follows a standard bivariate normal distribution with (unknown) correlation parameter  $\rho$ .

The vectors of individual observable characteristics,  $(Z_i, X_i)$  need not be the same across equations, and they will only include characteristics of the worker. Ideally, one would like to include characteristics of the various potential employers with whom a worker might be matched, but this information is not available in household surveys such as ours. Another piece of missing information is whether individuals who do not apply for a formal job would be hired in that position if they were to apply.<sup>26</sup>

In practice, for any given worker, there are three possible scenarios that can be distinguished in the data: i) being a formal worker, which occurs with probability  $P(V^a > 0, V^h > 0|Z, X)$ ; ii) being an involuntary informal worker, which occurs with probability  $P(V^a > 0, V^h \leq 0|Z, X)$ ; and iii) being a voluntary informal worker, which occurs with probability  $P(V^a \leq 0|Z, X)$ . These three scenarios are incorporated into a discrete choice model that captures the joint decisions of workers and formal sector employers.<sup>27</sup> The likelihood function of the discrete choice problem is

$$\mathcal{L} = \prod_{i \in \mathcal{V}} [1 - F(Z\gamma_a)] \prod_{i \in \mathcal{J}} G(Z\gamma_a, -X\gamma_h; -\rho) \prod_{i \in \mathcal{F}} G(Z\gamma_a, X\gamma_h; \rho) \quad (2)$$

where  $F(\cdot)$  and  $G(\cdot, \cdot; \cdot)$  are the standardized normal and bivariate normal distributions,  $\mathcal{V}$  is the set of voluntary informal workers,  $\mathcal{J}$  is the set of involuntary informal workers, and  $\mathcal{F}$  is the set of formal workers. This model can be estimated using maximum likelihood methods. The estimates of the model provide a reduced form *ceteris paribus* answer to the question of which individual characteristics are associated with applying for and being hired in a formal job.<sup>28</sup>

The discrete choice model can also be extended to estimate selectivity-corrected earnings equations for different types of workers through a switching-regression model that allows for sector allocation based on both the application and hiring decisions modelled in equations (1.1) and (1.2). In this case, it will be assumed that the error terms in the log-earnings equations and the

---

<sup>24</sup> In this context, ‘unobservable’ means characteristics not available to the econometrician.

<sup>25</sup> In the above equations and the ones that follow, I drop the individual subscript  $i$ .

<sup>26</sup> It is conceptually possible that voluntary informal workers could be offered a formal job *if they* were to apply for that job. However, we do not allow for that possibility in our econometric model.

<sup>27</sup> This model is sometimes called a bivariate probit with ‘sample selection’ (see for instance Van de Ven and Van Praag 1981).

<sup>28</sup> The model is a reduced form one because it does not explicitly incorporate wages in the choice decisions equations.

unobservables  $(u_{ai}, u_{hi})$  defined above are jointly distributed with a multivariate normal distribution. The selectivity-adjusted log-earnings functions will be

$$\log y_s = X_s \beta_s + \theta_{sa} \lambda_a(Z, X) + \theta_{sh} \lambda_h(Z, X) + \epsilon_s, \quad (3)$$

where the subscript  $s$  refers to the three different groups of workers characterized previously, i.e. formal, involuntary informal, and voluntary informal workers. The terms  $\lambda_a(Z, X)$  and  $\lambda_h(Z, X)$  are selectivity correction terms that adjust for the fact that individuals are not randomly assigned across sectors (Tunali 1986 presents a detailed exposition of this family of models).<sup>29</sup>

Through these selectivity-adjusted earnings equations, one can obtain a prediction of the counterfactual earnings that an informal worker would obtain if working in a formal job. This exercise can be performed separately for voluntary and involuntary informal workers.

Key demographic characteristics are included in the vector of observables  $Z$  that affect the probability of applying for formal employment. In particular,  $Z$  includes age (and its square), marital status, as well as the household dependency ratio, and the interaction of this dependency ratio with the number of adult females in the household. As previously mentioned, more dependents will need more hours of care and a higher income to sustain them. Since females typically do most of the housework, this variable is expected to negatively affect their probability of applying for a formal job as these jobs are less flexible in their schedules. However, as the number of adult females increases within a household, the housework load per woman will be smaller, hence the need to interact these two variables. Furthermore, the effects of these demographic variables are estimated separately by gender.

The vector  $Z$  also includes the respondents' years of schooling and a dummy indicating whether they are still enrolled in school. Finally, a set of dummy variables at the city level are included to control for varying conditions in local labour markets.

The variables included in vector  $X$ , which enters the hiring equation (1.2), are a subset of  $Z$ , and only include variables that would be observable by a formal sector employer, such as age, gender, marital status, the aforementioned schooling variables, and city-level dummies. In other words, the demographic characteristics of the household (i.e. the dependency ratio and its interaction with the number of adult females) are excluded, as these characteristics are typically not observable by employers and, thus are less likely to affect their hiring decisions. Finally, all three log-earnings equations (3) include age (and its square), years of tenure in the current job (and its square), gender and marital status (interacted), the above schooling variables, and city-level dummies.

The model is estimated over two samples, depending on the definition of what is considered a 'formal job'. One sample considers 'formal jobs' to be those jobs covered by social security benefits. As previously discussed, this essentially implies restricting the definition of 'formal jobs' to formal wage employment (see Table 1). All formal self-employed workers are therefore grouped with other informal workers.

In a second sample, the above definition is extended to include the formal self-employed in the set of formal jobs. In this alternative sample, the interpretation of the hiring equation (1.2) changes,

---

<sup>29</sup> In the case of the earnings equation of voluntary informal workers, the term  $\theta_{sh} \lambda_h(Z, X)$  is zero. A detailed presentation of the model is included in a companion online Appendix. A more detailed presentation, including the formulas for the standard errors of the parameters, can be found in Tunali (1986).

as it now describes factors that affect hiring by a formal sector employer, as well as the costs and benefits for a firm operating formally (see Ulyssea 2020 for a discussion of what these costs might be). While information is available from the Labour Force Survey about firm characteristics, these characteristics occur *after* the sector selection process has taken place, and thus are already an endogenous outcome of the job allocation process. It is not obvious how to best incorporate this information into the above econometric framework, and thus the vector  $X$  is left specified as before, i.e. with only individual sociodemographic characteristics of the worker. Therefore, for the second sample, equation (1.2) can be interpreted as a reduced form that captures the association of these sociodemographic characteristics with hiring and firm formalization decisions.<sup>30</sup>

## 5 Results

This section presents the results of the estimation of the models described above, beginning with the results pertaining to the discrete choice model.

### 5.1 Discrete choice model

The results of the parameter estimates of equations (1.1) and (1.2) are included in Tables A2 and A3 in the Appendix. Instead, this section presents the average partial effects as these are easier to interpret, i.e. we report the average derivatives of a probability of interest (e.g. the probability of applying for a formal job) with respect to one explanatory variable, holding all other observable variables constant.

Table 8 presents the partial effects for the sample where formality is determined by having a job with social security benefits. The first two columns of the table report the result of a standard probit which does not separate the applying from the hiring decisions. The last four columns report the parameter estimates of the discrete choice model in equations (1.1) and (1.2). The second pair of columns contains the corresponding partial effects of the ‘apply’ equation (1.1), while the last pair of columns presents the partial effects of the ‘hiring’ decision *conditional* on having applied for a formal job position, i.e. the average partial effects of covariates on the conditional probability:

$$\frac{P(V^a > 0, V^h > 0 | Z, X)}{P(V^a > 0 | Z, X)}$$

Analysing first the partial effects corresponding to the decision to apply to a formal job, columns 3 and 4 of Table 8 show a concave relation between the propensity to apply and age for females. For males, however, the partial effect of age is negative among young workers and statistically insignificant otherwise. Married workers are about 4 percentage points more likely to apply for formal jobs, although the effect is only statistically significant for males. Neither the education

---

<sup>30</sup> The second sample is smaller than the first one because I import information from the Labour Force Survey (ENOE) to determine who among the self-employed is formal, and in the process of merging both ENOE and MOTRAL, I discard observations with discrepancies at the job characteristics level (e.g. occupation, industry, etc.). Descriptive statistics for these samples are found in Table A1 in the Appendix.



variables nor the cumulative effect of being male have a statistically significant impact on the applying decision.

An increase of one unit in the dependency ratio decreases the probability of women applying for formal employment by almost 5 percentage points, but this negative effect disappears if there is another adult female present in the household, presumably because that other female will help with the care of the dependents. In contrast, the opposite effect is found for males, i.e. there is a positive association between this ratio and the propensity to apply. These findings support the idea that the division of responsibilities at home has an important influence on whether workers want formal jobs, but the effects differ depending on the gender of the worker. More specifically, having more dependents and no extra help leads women to search for more flexible informal jobs, as these will allow them to care for dependents. For men, in contrast, having more dependents increases the propensity to apply for formal jobs.

Regarding the probability of being hired in a formal job, conditional on having applied, this probability presents an inverted-u association with age for females and a u-shaped association for males. Married men are almost 12 percentage points more likely to be hired in formal employment than single males, while the opposite is the case for married females. This hiring penalty for married females likely reflects that formal sector employers perceive them as less attractive employees, a finding that is consistent with recent experimental evidence for Mexico (Arceo-Gomez and Campos-Vazquez 2014). These findings, together with those pertaining to the applying decision, highlight the importance of the division of labour at the household level in shaping both applying and hiring decisions in the formal sector.

Finally, education is a key factor affecting the probability of being hired. In particular, an extra year of schooling increases this probability by about 3 percentage points, but being enrolled in school decreases this hiring probability by almost 25 percentage points.<sup>31</sup>

Note also that the partial effects of the standard probit model (in the first two columns of Table 8) are a mix of the effects identified by the bivariate probit model, which separates applying from hiring decisions. However, without the bivariate probit, it would not be possible to disentangle which factors affect each of these separate decisions.

A comparison of the above results with those in Table 9, where a formal job includes jobs with coverage as well as the formal self-employed, shows that the estimated partial effects are very similar and the differences are mainly in the magnitudes and statistical significance of some effects. Among these differences, one can note that marital status has a weaker effect on the hiring equation in this sample. Also, the effects of the dependency ratio on the application decision for men are stronger, and the years of schooling positively affect the probability of applying for a formal job, while in the previous sample this effect was statistically insignificant.

## 5.2 Selectivity-adjusted earnings equations

If the above discrete choice model is complemented with log-earnings equations, one can obtain parameters for the latter which are adjusted for potential sample-selectivity biases. As not all

---

<sup>31</sup> Appendix Table A2 reports the correlation parameter between the unobservables ( $u_a$ ,  $u_h$ ), which is -0.53 (and statistically significant at the 1 per cent level), indicating that the unobservable factors that lead to a greater probability of applying for formal employment are negatively correlated with the unobservables that affect the probability of being hired.

individuals report their earnings or have positive earnings, the previous discrete choice models are re-estimated for a subsample of workers with positive earnings and are used to estimate the selectivity correction terms in equation (3).<sup>32</sup> The results of these estimations are presented in Tables 10 and 11.

Table 10 shows that, consistent with standard Mincerian equations, age has increasing concave returns, with inflection points around ages 37–47. However, this shape is statistically significant only among formal workers.

In general, males obtain substantially higher earnings relative to otherwise comparable single females. The earnings premiums range between 20 per cent for single males in formal employment and 60 per cent for married males in voluntary informal employment.<sup>33</sup> In contrast, among voluntary informal workers, married women display a *ceteris paribus* earnings penalty of more than 30 per cent relative to their single female counterparts. These numbers indicate that demographic factors at the household level affect the job allocation process and the earnings obtained in the market.

The number of years of tenure at a firm have increasing concave returns. However, this shape is statistically significant only among involuntary informal workers, with an inflection point around 14 years of tenure. An extra year of schooling is associated with an increase in earnings of about 15 per cent in formal jobs and 8 per cent among those who do not want a job with coverage, and has no significant effect on the earnings of involuntary informal workers. Being enrolled in school leads to *ceteris paribus* earnings losses of almost 40 per cent among formal workers only.

Finally, most of the coefficients for the selectivity correction terms are statistically insignificant at the 90 per cent level. The exception occurs for the correction term in the earnings equation of voluntary informal workers. In this case, the unobserved factors that affect the probability of applying for formal jobs are negatively correlated with the unobservables in the earnings equation of these workers.

It is worth mentioning that the fit of the earnings equations is higher for workers employed in their preferred sector, i.e. formal and voluntary informal workers. This indicates that there is a greater degree of unexplained heterogeneity in the earnings of involuntary informal workers.

The corresponding estimations for the sample which includes the formal self-employed among formal workers are reported in Table 11. The results are qualitatively similar to those of the other sample and, for the sake of brevity, I omit further discussion of them.

Using these regression models, counterfactual earnings of informal workers are predicted if they worked in formal employment. These differentials are presented in Appendix Tables A4 and A5. None of these differentials are statistically different to zero and they are quite sensitive to the

---

<sup>32</sup> The parameter estimates of the bivariate probit over this slightly more restricted sample are available upon request from the author.

<sup>33</sup> These premia are obtained by noting that for a given gender  $x$  marital group  $g$ , the parameters estimated are  $\hat{\beta}_g = \ln(\widehat{y}_g) - \ln(\widehat{y}_{sf})$ , where  $\ln(\widehat{y}_g)$  and  $\ln(\widehat{y}_{sf})$  are the predicted log earnings of group  $g$  and single females, respectively. Hence  $\exp(\hat{\beta}_g) - 1$  approximates the percentage premium for group  $g$  relative to single females.

selected functional forms chosen for the estimation of the econometric models in equations (1.1), (1.2), and (3). Hence, these results should be interpreted with caution.

In summary, the econometric analysis performed in this section indicates that the division of labour at home across gender lines plays an important role in determining who applies for and who is hired in formal job vacancies, as well as the earnings gained in formal employment. Also, the levels of schooling are a crucial factor affecting both the probability of being hired in a formal job and the earnings obtained in it.

## 6 Conclusions

This paper exploits a unique dataset containing information about the preferred type of jobs of workers in large urban centres in Mexico. Comparing this information with the actual jobs they have, it is estimated that almost 80 per cent of informal workers consider it preferable to be employed in a formal job which provides them with social security coverage, even if it entails paying the corresponding taxes for such benefits. This suggests that many of the urban informal workers are in this sector because of a lack of better options.

Our econometric analysis highlights two important factors that affect the workers' access to formal wage employment, i.e. the division of housework at the household level and the levels of human capital. In particular, the traditional division of labour at home is a likely culprit for limiting the willingness of females to apply for formal wage jobs and the probability of being hired in such jobs. Also, having a higher level of education plays a significant role in increasing the chances of being hired in a formal job and of earning a higher income from it. While one should not draw direct policy recommendations from these findings, it seems important to consider how policy can affect these two dimensions in order to encourage the successful transition of workers into formal employment.

Finally, one methodological point arises from this research. So far, the overwhelming majority of research on labour markets in developing countries is based on variables such as wages, employment status, and so forth. The analysis conducted here shows that there is much to be gained by also considering the stated preferences of workers about potential jobs and their characteristics.

This new piece of information can enrich our understanding of the functioning of labour markets and the welfare of workers. In particular, this type of information can help to solve some unresolved theoretical debates in the literature, where traditional analyses have led to ambiguous conclusions. To exploit this type of information, however, will require a better data collection effort, including carefully worded questions for eliciting workers' preferences for different types of job characteristics.

## References

Alaniz, E., T.H. Gindling, C. Mata, and D. Rojas (2020). 'Transforming Informal Work and Livelihoods in Costa Rica and Nicaragua'. WIDER Working Paper 100/2020. Helsinki: UNU-WIDER. <https://doi.org/10.35188/UNU-WIDER/2020/857-3>

- Alcaraz, C., D. Chiquiar, and A. Saucedo (2015). 'Informality and Segmentation in the Mexican Labor Market'. Banco de México Working Papers 2015-25. Mexico City: Mexico. <https://doi.org/10.36095/banxico/di.2015.25>
- Arceo-Gomez, E.O., and R.M. Campos-Vazquez (2014). 'Race and Marriage in the Labor Market: A Discrimination Correspondence Study in a Developing Country'. *American Economic Review*, 104(5): 376–80. <https://doi.org/10.1257/aer.104.5.376>
- Bosch, M., and W.F. Maloney (2010). 'Comparative Analysis of Labor Market Dynamics Using Markov Processes: An Application to Informality'. *Labour Economics*, 17(4): 621–31. <https://doi.org/10.1016/j.labeco.2010.01.005>
- Contreras, D., R. Gillmore, and E. Puentes (2017). 'Self-Employment and Queues for Wage Work: Evidence from Chile'. *Journal of International Development*, 24(4): 473–99. <https://doi.org/10.1002/jid.3074>
- Duval-Hernández, R., and R. Smith (2010). 'Counterfactual Wages and Segmentation between Formal and Informal Sectors in Mexico'. Unpublished Manuscript.
- Fields, G.S. (1975). 'Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs'. *Journal of Development Economics*, 2(2): 165–87. [https://doi.org/10.1016/0304-3878\(75\)90014-0](https://doi.org/10.1016/0304-3878(75)90014-0)
- Fields, G.S. (1990) 'Labour Market Modelling and the Urban Informal Sector: Theory and Evidence'. In D. Turnham, B. Salome, and A. Schwarz (eds), *The Informal Sector Revisited*. Paris: Organisation for Economic Co-operation and Development.
- Fields, G.S. (2011). "Informality": It's Time to Stop Being Alice-in-Wonderland-ish'. WIEGO Research Paper. Manchester: WIEGO.
- Fields, G.S. (forthcoming). 'Informality and Work Status'. In G.S. Fields and T. Gindling, *Transforming Informal Work and Livelihoods*.
- Fiess, N.M., M. Fugazza, and W.F. Maloney (2010). 'Informal Self-employment and Macroeconomic Fluctuations'. *Journal of Development Economics*, 91(2): 211–26. <https://doi.org/10.1016/j.jdeveco.2009.09.009>
- Gindling, T.H., and D. Newhouse (2014). 'Self-Employment in the Developing World'. *World Development*, 56: 313–31. <https://doi.org/10.1016/j.worlddev.2013.03.003>
- Günther, I., and A. Launov (2012). 'Informal Employment in Developing Countries: Opportunity or Last Resort?'. *Journal of Development Economics*, 97: 88–98.
- INEGI (Instituto Nacional de Estadística y Geografía) (2012). 'Encuesta Nacional de Micronegocios'. Available at: <https://www.inegi.org.mx/programas/enamin/2012/> (accessed 13 September 2020).
- INEGI (Instituto Nacional de Estadística y Geografía) (2014). 'La informalidad laboral: Encuesta Nacional de Ocupación y Empleo: marco conceptual y metodológico'. México: INEGI.
- INEGI (Instituto Nacional de Estadística y Geografía) (2015a). 'Módulo de Trayectorias Laborales (MOTRAL) 2015'. Available at: <https://www.inegi.org.mx/programas/motral/2015/> (accessed 1 July 2019).
- INEGI (Instituto Nacional de Estadística y Geografía) (2015b). 'Encuesta Nacional de Ocupación y Empleo, población de 15 años y más de edad'. Available at: <https://www.inegi.org.mx/programas/enoe/15ymas/> (accessed 1 November 2018).
- Levy, S. (2008). *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Washington, DC: Brookings Institution Press.
- Levy, S. (2018). *Under-Rewarded Efforts: The Elusive Quest for Prosperity in Mexico*. Washington: Inter-American Development Bank. <https://doi.org/10.18235/0001189>

- Magnac, T. (1991). 'Segmented or Competitive Labor Markets'. *Econometrica*, 59(1): 165–87. <https://doi.org/10.2307/2938245>
- Maloney, W.F. (1999). 'Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico'. *World Bank Economic Review*, 13(2): 275–302. <https://doi.org/10.1093/wber/13.2.275>
- Maloney, W.F. (2004). 'Informality Revisited'. *World Development*, 32(7): 1159–78. <https://doi.org/10.1016/j.worlddev.2004.01.008>
- Marcouiller, D., V.R. de Castilla, and C. Woodruff (1997). 'Formal Measures of the Informal-Sector Wage Gap in Mexico, El Salvador, and Peru'. *Economic Development and Cultural Change*, 45(2): 367–92. <https://doi.org/10.1086/452278>
- Ruppert Bulmer, E. (2018) 'Defining Informality vs Mitigating Its Negative Effects'. *IZA World of Labor*, 442. <https://doi.org/10.15185/izawol.442>
- Soares, F.V. (2004). 'Do Informal Workers Queue for Formal Jobs in Brazil?'. Texto para Discussão #1021, Rio de Janeiro: IPEA.
- Tunali, İ. (1986). 'A General Structure for Models of Double Selection and an Application to a Joint Migration/Earnings Process with Remigration'. *Research in Labor Economics*, 8: 235–82.
- Ulysea, G. (2011). 'The Formal-Informal Labor Market Segmentation Hypothesis Revisited'. *Brazilian Review of Econometrics*, 30(2): 311. <https://doi.org/10.12660/bre.v30n22010.3467>
- Ulysea, G. (2020). 'Informality: Causes and Consequences for Development'. *Annual Review of Economics*, 12(1): 525–46. <https://doi.org/10.1146/annurev-economics-082119-121914>
- Van de Ven, W.P.M.M., and B.M.S. Van Praag (1981). 'The Demand for Deductibles in Private Health Insurance. A Probit Model with Sample Selection'. *Journal of Econometrics*, 17(2): 229–52. [https://doi.org/10.1016/0304-4076\(81\)90028-2](https://doi.org/10.1016/0304-4076(81)90028-2)

## Tables and figures

Table 1: Work status definition and operationalization, the case of Mexico

Work status group	Definition/operationalization
Formal self-employed	Formal self-employed workers are the self-employed (own-account or owners) who operate a business registered with tax authorities and with fixed work premises. Also, all employers (with at least one employee) in the agricultural sector enter this category.
Upper-tier informal self-employed	Upper-tier informal self-employed (own-account and owners) are those who voluntarily enrol for social security coverage (through the government or have private insurance), or work in a profession that requires post-secondary or vocational education, or they are employers with at least one employee, or if their place of work has fixed premises.
Lower-tier informal self-employed	All other self-employed not in the above two categories. This includes all self-employed working in agriculture.
Formal wage employees	Wage employees whose employers contribute to social security, except if the employer itself is a non-registered business (i.e. the firm is not registered with tax authorities and does not have fixed work premises).
Upper-tier informal wage employees	Wage employees whose employers do not contribute to social security (or if they do, the employer itself is a non-registered business) BUT who receive other benefits such as paid annual leave, profit-sharing, (government-sponsored) housing credit, day-care facilities, private insurance (life or health), saving funds, time for parental care, or <i>aguinaldo</i> (mandatory one-month salary bonus in December), OR work in a profession that requires post-secondary or vocational education, OR they have a permanent contract.
Lower-tier informal wage employees	Lower-tier informal employees are all other employees. Also, all unpaid workers are included in this category.

Source: author's illustration.

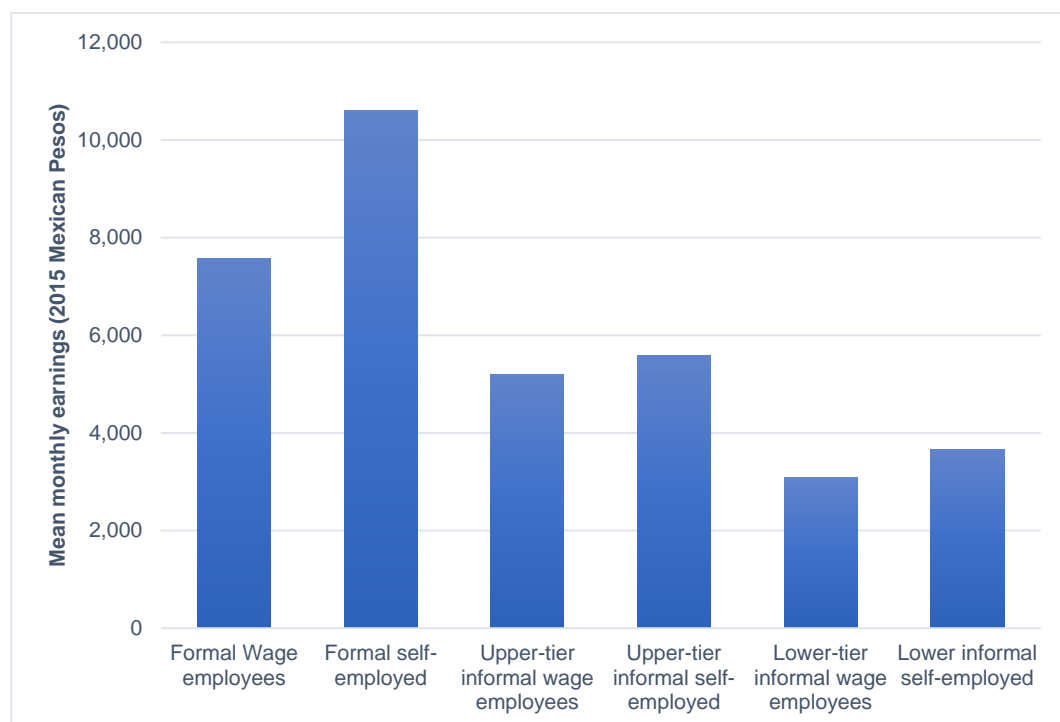
Table 2: Descriptive statistics of Mexican labour markets, 2015, large urban centres

	Formal employment		Upper-tier Informal		Lower-tier Informal		Unemp.	Out of labour force
	Wage empl.	Self-empl.	Wage empl.	Self-empl.	Wage empl.	Self-empl.		
<b>Percentage*</b>	<b>46.9</b>	<b>5.8</b>	<b>10.1</b>	<b>5.6</b>	<b>20.3</b>	<b>11.3</b>	<b>5.2</b>	<b>39.0</b>
Years of age	37.9	46.6	36.2	45.1	36.6	46.6	32.5	41.6
Male (%)	60.2	71.7	54.2	67.3	55.0	54.2	56.7	29.0
Years of schooling	12.3	13.2	11.1	10.1	8.7	8.2	11.4	9.3
Schooling level (%)								
Elementary	8.6	10.5	16.9	25.8	30.7	38.0	10.8	26.82
Intermediate	54.4	37.2	52.8	53.3	61.8	55.4	60.3	56.1
Higher	37.0	52.4	30.2	20.9	7.5	6.6	28.9	17.08
Enrolled in school (%)	5.2	2.2	10.0	3.0	10.0	2.1	10.8	29.5
Married (%)	46.9	64.7	34.3	54.4	36.1	50.3	24.0	39.8
Household comp.								
Dependency ratio	0.56	0.52	0.56	0.60	0.67	0.67	0.53	0.71
# Adult females	1.21	1.18	1.18	1.18	1.20	1.06	1.28	0.95
Occupation (%)								
Managers	7.5	12.1	3.4	1.5	0.1	0.0		
Professionals	28.7	34.4	34.9	35.0	1.4	0.0		
Clerical	15.6	0.7	12.1	1.4	2.3	0.1		
Sales	10.5	29.7	10.0	26.3	19.7	27.8		
Services	8.5	8.0	8.2	12.6	11.9	9.8		
Manual	4.2	10.0	7.3	20.0	13.4	28.6		
Operators	15.0	3.5	6.4	0.9	11.1	7.5		
Elementary	10.1	1.6	17.7	2.4	40.2	26.2		
Industry (%)								
Primary	1.4	1.3	0.7	0.0	1.7	2.4		
Construction	4.4	4.6	5.8	12.4	11.9	11.1		
Manufacturing	23.2	8.4	10.8	10.1	13.2	9.7		
Trade	15.9	31.2	11.5	26.9	22.1	44.0		
Communication	6.1	4.9	3.7	0.7	9.2	8.1		
Business services	13.5	21.9	16.4	10.3	3.5	3.5		
Education	10.5	1.5	7.6	4.1	0.6	0.0		
Health	6.1	7.0	6.1	3.0	0.8	0.0		
Personal services	8.3	19.2	32.3	32.6	36.2	21.3		
Public admin.	10.7	0.0	5.1	0.0	0.7	0.0		
Public sector (%)	24.9	0.0	9.9	0.0	1.7	0.0		
Social security (%)	100	0.1	0.0	0.0	0.0	0.0		
Earnings	7,576	10,598	5,201	5,589	3,085	3,667		

Note: the definitions of the different work statuses follow the criteria in Table 1. The percentages are the shares of total employment in columns 1–6, the unemployment rate in column 7, and the population out of the labour force as a percentage of the total working-age population in column 8. The variable # of adult females counts the number of females aged 18+ in the household, excluding the survey respondent in the case of female respondents. Earnings are measured in 2015 Mexican Pesos per month and include individuals with zero earnings. All estimates use sampling weights.

Source: author's calculations based on ENOE 2<sup>nd</sup> Quarter 2015, large urban areas for individuals aged 15 years and more (INEGI 2015b).

Figure 1: Mean earnings by work status



Note: all estimates use sampling weights.

Source: author's calculations based on ENOE 2<sup>nd</sup> Quarter 2015, urban areas (INEGI 2015b).

Table 3: One-year transitions across work status

		1st Quarter 2016							Share of stayers	
		Self-employed			Wage-employed					
		Formal	Informal		Formal	Informal				
			Upper	Lower		Upper	Lower			
1st Quarter 2015	Self-employed	Formal		52.73	15.28	7.11	10.13	6.04	8.71	3.54
		Informal	Upper	18.47	35.26	16.99	5.81	6.77	16.7	1.94
			Lower	5.78	8.68	59.2	6.5	2.21	17.64	5.97
	Wage-employed	Formal		0.96	0.87	1.63	87.44	5.11	3.99	44.42
		Informal	Upper	2.94	3.57	2.83	27.84	33.92	28.89	3.2
			Lower	4.5	3.99	9.54	12.84	12.02	57.11	9.96
TOTAL		6.69	5.32	10.15	50.94	8.9	18	69.03		

Note: all estimates use sampling weights.

Source: author's calculations based on ENOE 1<sup>st</sup> Quarter 2015 and 2016, urban areas (INEGI 2015b).



Table 4: Preference for jobs with social security

	Prefers job with social security [1]	Most-valued attribute of social security				
		Health insurance	Pension	Life insurance	Housing benefits	Disability benefits
		[2]	[3]	[4]	[5]	[6]
<b>Formal employment</b>						
Wage employment	83.6	46.2	22.8	13.2	13.7	4.2
Self-employment	77.7	44.5	25.2	16.4	6.8	7.1
<b>Informal employment</b>						
Upper-tier						
Wage employment	73.7	55.2	15.3	11.5	10.2	7.8
Self-employment	75.3	52.2	14.3	4.5	25.3	3.7
Lower-tier						
Wage employment	79.3	41.3	14.1	21.3	16.4	6.9
Self-employment	86.3	39.4	19.8	13.6	19.4	7.8
Unemployed	92.6	49.2	27.2	14.3	5.0	4.3
Out of labour force	85.0	52.1	17.0	10.9	13.4	6.6

Note: column 1 reports the percentage of individuals who consider it is better to have a job with social security benefits, even if one must pay to be entitled to such benefits. Columns 2–5 show the most-valued reason for preferring social security. The numbers in these columns are the row percentages of respondents selecting a given reason among those individuals who declare preferring a job with social security coverage. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table 5: Characteristics of employed population in 2015 MOTRAL module

	Has social security		Does not have social security	
	Wants social security	Does not want social security	Wants social security	Does not want social security
	[1]	[2]	[3]	[4]
Years of age	37.1	35.4	38.4	37.1
Male (%)	48.2	63.0	48.4	49.7
Years of schooling	11.8	12.4	9.9	10.7
Enrolled in school (%)	3.9	2.9	6.1	6.3
Married (%)	45.1	42.5	44.7	32.7
Household composition				
Dependency ratio	0.61	0.59	0.61	0.72
# Adult females	1.06	1.55	1.06	1.00
Earnings (monthly Mexican Pesos)	7,004	9,075	4,320	3,599
Share of employment	43.5	8.9	38.3	9.3
# Obs.(unweighted)	1,906	393	1,759	439

Note: unless otherwise stated, all numbers are averages across the different employment groups. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table 6: One-quarter transitions across work status and preference for employment with social security coverage

			2nd Quarter 2015										Number of inweighted obs.
			Self-employed			Wage-employed			Unemployment	Out of labour force	Share of stayers		
			Formal	Informal		Formal	Informal						
				Upper	Lower		Upper	Lower					
1st Quarter 2015	Self-employed	Formal	48.42	10.38	4.18	18.09	4.26	2.55	6.39	5.72	1.63	162	
		Informal	Upper	32.67	26.03	5.47	3.01	3.71	11.22	1.74	16.16	1	151
			Lower	2.36	7.76	58.45	2.8	0.9	12.22	1.15	14.35	4.66	317
	Wage-employed	Formal	1.11	0.25	0.46	89.14	2.9	2.26	2.61	1.25	34.67	1796	
		Informal	Upper	4.18	3.83	0.98	29.81	36.94	13.78	4.55	5.93	2.76	383
			Lower	1.55	0.82	8.48	11.37	8.84	52.3	6.01	10.63	8	655
	Unemployment		4.26	2.16	5.34	31.77	3.3	12.36	17.54	23.26	0.61	197	
	Out of the labour force		1.01	1.71	5.92	6.93	2.54	10.92	4.25	66.71	13.17	944	
TOTAL		4.39	2.89	7.91	42.04	6.21	13.98	4.09	18.48	66.5	4605		
			2nd Quarter 2015										Number of unweighted obs.
			Self-employed			Wage-employed			Unemployment	Out of labour force	Share of stayers		
			Formal	Informal		Formal	Informal						
				Upper	Lower		Upper	Lower					
			1st Quarter 2015	Self-employed	Formal	47.94	20.69	9.8	4.14	1.2	12.75	0	3.48
Informal	Upper	24.44			38.82	5.52	18.04	3.06	6.24	0	3.89	1.32	43
	Lower	1.59			6.72	40.39	0.77	0	37.76	0	12.78	3.2	83
Wage-employed	Formal	1.51		0.4	1.31	87.31	5.23	1.22	0.56	2.47	31.24	373	
	Informal	Upper		27.75	10.31	0.87	24.99	16.44	11.38	3.06	5.21	1.71	78
		Lower		0.29	2.01	1.39	17.41	17.38	54.45	0.81	6.27	8.26	150
Unemployment		0.83		0	3.52	6.96	22.72	14.97	6.03	44.98	0.54	61	
Out of the labour force		0.48		3.34	7.63	6.91	13.36	16.05	2.36	49.87	7.76	173	
TOTAL		5.95	4.48	5.94	38.97	10.46	17.28	1.54	15.38	55.4	1005		

Note: transitions for individuals who consider it is better to not have a job with social security benefits. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 1<sup>st</sup> and 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table 7: Preferences for wage employment among micro-entrepreneurs in 2012

	Accept wage job with social security at going wage %	Will not quit self- employment %	%	Other self-employed	
				Median reservation earnings for wage job	
				2012 Mx Pesos	Multiple of median formal earnings
	[1]	[2]	[3]	[4]	[5]
Formal self-employment	46.2	43.3	10.5	20,000	3.8
Informal self-employment					
Upper-tier	56.4	37.2	6.3	14,000	2.7
Lower-tier	64.5	31.8	3.6	8,000	1.5
Total	59.1	35.3	5.6	12,000	2.3

Note: estimates for micro-entrepreneurs in large urban areas aged 18–54 years and who had more than one year operating their business. All estimates use sampling weights.

Source: author's calculations based on ENAMIN 2012 (INEGI 2012).

Table 8: Average partial effects of discrete choice models. Formality status defined by job with social security

	Probit			Bivariate probit with selectivity					
	[1]	[2]		Apply	[4]		Hire   Apply	[6]	
Age									
Females at 18 years	0.019	(0.004)	***	0.0083	(0.003)	**	0.0170	(0.004)	***
Females at 35 years	0.0022	(0.001)	***	0.0014	(0.001)	**	0.0014	(0.001)	*
Females at 54 years	-0.017	(0.005)	***	-0.003	(0.003)		-0.015	(0.005)	***
Males at 18 years	-0.017	(0.007)	**	-0.003	(0.001)	***	-0.013	(0.004)	***
Males at 35 years	-0.008	(0.002)	***	-0.0031	(0.002)		-0.0066	(0.002)	***
Males at 54 years	0.006	(0.008)		0.0013	(0.005)		0.0048	(0.006)	
Married									
Females	-0.084	(0.023)	***	0.042	(0.027)		-0.130	(0.013)	***
Males	0.124	(0.016)	***	0.048	(0.022)	**	0.116	(0.022)	***
Male	0.0006	(0.018)		-0.0135	(0.009)		0.0169	(0.013)	
Years of schooling	0.029	(0.003)	***	0.0010	(0.003)		0.031	(0.003)	***
Enrolled in school	-0.22	(0.073)	***	-0.030	(0.035)		-0.24	(0.084)	***
Dependency ratio (DR)									
Females									
DR w/ no adult females	-0.066	(0.027)	**	-0.049	(0.025)	*			
DR w/ 1 adult female	0.021	(0.020)		0.0035	(0.008)				
Males									
DR w/ no adult females	0.063	(0.031)	*	0.025	(0.012)	**			
DR w/ 1 adult female	0.049	(0.046)		0.016	(0.010)				

Note: standard errors robust to clustering at the city level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Inference done with a t-distribution with # clusters -1 as degrees of freedom. The variable # of adult females counts the number of females aged 18+ in the household, excluding the survey respondent in the case of female respondents. City fixed effects included. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table 9: Average partial effects of discrete choice models. Formality status defined by job with social security or formal self-employment

	Probit			Bivariate probit with selectivity					
	[1]	[2]		Apply		Hire   Apply			
				[3]	[4]		[5]	[6]	
<b>Age</b>									
Females at 18 years	0.015	(0.004)	***	0.0097	(0.003)	***	0.0120	(0.004)	***
Females at 35 years	0.0054	(0.001)	***	0.0026	(0.001)	***	0.0037	(0.001)	***
Females at 54 years	-0.005	(0.006)	***	-0.001	(0.002)		-0.004	(0.005)	
Males at 18 years	-0.019	(0.008)	**	-0.004	(0.001)	***	-0.015	(0.007)	**
Males at 35 years	-0.003	(0.002)		-0.0021	(0.002)		-0.0018	(0.001)	
Males at 54 years	0.018	(0.009)	*	0.0045	(0.004)		0.0144	(0.007)	*
<b>Married</b>									
Females	-0.042	(0.037)		0.040	(0.022)	*	-0.080	(0.028)	***
Males	0.120	(0.030)	***	0.060	(0.028)	**	0.091	(0.048)	*
Male	0.0047	(0.021)		-0.0152	(0.011)		0.0242	(0.015)	
Years of schooling	0.045	(0.003)	***	0.0069	(0.002)	***	0.043	(0.004)	***
Enrolled in school	-0.19	(0.054)	***	-0.031	(0.025)		-0.20	(0.061)	***
<b>Dependency ratio (DR)</b>									
<b>Females</b>									
DR w/ no adult females	-0.046	(0.014)	***	-0.044	(0.020)	**			
DR w/ 1 adult female	0.032	(0.026)		0.0123	(0.015)				
<b>Males</b>									
DR w/ no adult females	0.085	(0.037)	**	0.083	(0.026)	***			
DR w/ 1 adult female	0.089	(0.051)	*	0.069	(0.028)	**			

Note: standard errors robust to clustering at the city level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Inference done with a t-distribution with # clusters -1 as degrees of freedom. The variable # of adult females counts the number of females aged 18+ in the household, excluding the survey respondent in the case of female respondents. City fixed effects included. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table 10: Selectivity-adjusted log-earnings OLS equations. Formality status defined by job with social security

	Has social security	Does not have social security	
		Wants social security	Does not want social security
Age	0.0271** (0.0106)	0.0425 (0.0322)	0.0516 (0.0419)
Age sq.	-0.000367*** (0.000119)	-0.000444 (0.000318)	-0.000610 (0.000561)
Single female (omitted)			
Married female	-0.0540 (0.0891)	-0.145 (0.152)	-0.384** (0.161)
Single male	0.182** (0.0679)	0.439*** (0.123)	0.401*** (0.100)
Married male	0.405*** (0.0462)	0.307*** (0.104)	0.472*** (0.101)
Tenure	0.0392 (0.0328)	0.0245*** (0.00643)	0.00984 (0.0371)
Tenure sq.	-0.00144 (0.00133)	-0.00104** (0.000387)	-0.0000236 (0.00118)
Years schooling	0.147*** (0.0360)	0.0391 (0.0385)	0.0792*** (0.0194)
Enrolled in school	-0.467** (0.204)	-0.130 (0.345)	-0.315 (0.191)
$\lambda(\text{apply})$	-0.569 (0.462)	-0.363 (0.669)	-0.403** (0.183)
$\lambda(\text{hire})$	0.788 (0.510)	-0.280 (0.836)	
Constant	5.756*** (0.596)	6.346*** (1.150)	4.853*** (0.667)
$R^2$	0.336	0.158	0.487
$N$	2091	1535	364

Note: standard errors robust to clustering at the city level and adjusted for generated regressors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Inference done with a t-distribution with # clusters -1 as degrees of freedom. Formal workers are those in a job with social security coverage. City fixed effects included. Sample includes only individuals with positive earnings. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table 11: Selectivity-adjusted log-earnings OLS equations. Formality status defined by job with social security or formal self-employment

	Formal employment	Informal employment	
		Wants social security	Does not want social security
Age	0.000953 (0.0168)	0.0726 (0.0640)	0.0987 (0.0635)
Age sq.	-0.0000264 (0.000193)	-0.000859 (0.000766)	-0.00140 (0.000874)
Single female (omitted)			
Married female	0.100 (0.0882)	-0.296** (0.110)	-0.340** (0.135)
Single male	0.172** (0.0636)	0.501*** (0.156)	0.545*** (0.131)
Married male	0.471*** (0.0731)	0.375*** (0.0763)	0.390*** (0.103)
Tenure	0.0372 (0.0354)	0.0120 (0.00976)	-0.000687 (0.0341)
Tenure sq.	-0.00125 (0.00128)	-0.00121** (0.000495)	-0.0000208 (0.000953)
Years schooling	0.157*** (0.0337)	0.0401 (0.0474)	0.0357** (0.0159)
Enrolled in school	-0.659 (0.397)	-0.114 (0.445)	-0.397** (0.169)
$\lambda(\text{apply})$	-0.337 (0.345)	-0.115 (0.226)	-0.339** (0.151)
$\lambda(\text{hire})$	0.979** (0.471)	0.187 (0.453)	
Constant	5.979*** (0.555)	6.165*** (1.139)	4.767*** (0.943)
$R^2$	0.293	0.193	0.472
N	2026	1131	276

Note: standard errors robust to clustering at the city level and adjusted for generated regressors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Inference done with a t-distribution with # clusters -1 as degrees of freedom. Formal workers are those in a job with social security coverage and the formal self-employed (see Table 1 for details). City fixed effects included. Sample includes only individuals with positive earnings. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).



## Appendix

Table A1: Descriptive statistics MOTRAL samples

	Sample 1			Sample 2		
	Has social security	Does not have social security Wants social security	Does not have social security Does not want social security	Formal work	Informal employment Wants social security	Does not want social security
<b>Percentage</b>	<b>52.4</b>	<b>38.3</b>	<b>9.3</b>	<b>59.7</b>	<b>32.1</b>	<b>8.2</b>
Years of age	36.8	38.4	37.1	37.8	37.8	35.6
Male (%)	50.7	48.4	49.7	49.3	44.8	48.3
Years of schooling	11.9	9.9	10.7	12.1	9.1	9.8
Schooling level (%)						
Elementary	10.2	24.2	16.7	10.7	27.6	19.4
Intermediate	56.1	58.9	61.0	54.1	63.0	69.0
Higher	33.6	16.9	22.2	35.3	9.5	11.6
Enrolled in school (%)	3.7	6.1	6.3	3.8	5.3	5.9
Married (%)	44.7	44.7	32.7	46.7	42.2	26.6
Household comp.						
Dependency ratio	0.60	0.61	0.72	0.60	0.62	0.71
# Adult females	1.14	1.06	1.00	1.13	1.01	0.96
Occupation (%)						
Managers	9.2	1.3	1.2	9.2	0.7	0.6
Professionals	26.9	18.0	20.4	29.0	12.4	13.5
Clerical	18.0	2.4	2.6	15.6	2.2	1.4
Sales	10.1	19.3	26.2	11.9	16.5	29.5
Services	8.5	13.7	6.7	8.1	16.2	6.5
Manual	5.0	12.7	18.8	5.0	12.7	19.0
Operators	12.2	8.6	6.6	11.6	10.5	8.1
Elementary	10.1	24.0	17.6	9.7	28.9	21.4
Industry (%)						
Primary	1.1	1.0	2.1	1.2	0.5	2.1
Construction	6.9	8.0	9.7	6.2	7.6	10.2
Manufacturing	21.3	10.9	14.7	19.0	13.5	16.5
Trade	15.3	22.8	26.2	16.7	20.8	30.3
Communication	5.4	6.5	5.7	4.8	7.9	6.4
Business services	15.1	10.0	11.8	17.2	6.6	3.6
Education	10.8	1.6	4.6	10.4	1.8	4.2
Health	4.4	3.1	0.3	4.2	1.4	0.3
Personal services	10.0	34.8	23.6	11.2	39.0	25.5
Public admin.	9.8	1.2	1.3	9.1	1.0	1.1
Earnings	7,344	4,320	3,599	7,363	3,517	2,853
Number of obs.	2,299	1,759	439	2,258	1,278	333

Note: the variable # of adult females counts the number of females aged 18+ in the household, excluding the survey respondent in the case of female respondents. Earnings are measured in 2015 Mexican Pesos per month. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table A2: Parameter estimates of discrete choice models. Formality status defined by job with social security

	Bivariate probit with selectivity		
	Probit	Apply	Hire
Females			
Age	0.106*** (0.0251)	0.0658** (0.0253)	0.0756*** (0.0191)
Age sq.	-0.00142*** (0.000368)	-0.000796** (0.000364)	-0.00104*** (0.000277)
Males			
Age	-0.0908 (0.0576)	-0.0647 (0.0740)	-0.0642** (0.0289)
Age sq.	0.000987 (0.000744)	0.000663 (0.000907)	0.000702* (0.000378)
Single female (omitted)			
Married female	-0.236*** (0.0660)	0.287 (0.203)	-0.390*** (0.0374)
Single male	3.397*** (1.077)	2.488** (1.205)	2.360*** (0.634)
Married male	3.738*** (1.059)	2.785** (1.335)	2.634*** (0.608)
Females			
Dependency ratio	-0.184** (0.0736)	-0.272** (0.130)	
Dependency ratio x # Adult females	0.241** (0.0925)	0.298** (0.139)	
Males			
Dependency ratio	0.173* (0.0855)	0.169** (0.0815)	
Dependency ratio x # Adult females	-0.0366 (0.0796)	-0.0684 (0.0617)	
Years schooling	0.0817*** (0.0103)	0.00664 (0.0205)	0.0846*** (0.00805)
Enrolled in school	-0.637*** (0.224)	-0.172 (0.185)	-0.626** (0.238)
Constant	-2.542*** (0.340)	0.0815 (0.517)	-1.786*** (0.277)
Rho			-0.534*** (0.154)
Number of obs.	4,497		4,497

Note: standard errors robust to clustering at the city level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The variable # of adult females counts the number of females aged 18+ in the household, excluding the survey respondent in the case of female respondents. City fixed effects included. Inference done with a t-distribution with # clusters -1 as degrees of freedom. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table A3: Parameter estimates of discrete choice models. Formality status defined by job with social security or formal self-employment

	Probit	Bivariate probit with selectivity	
		Apply	Hire
Females			
Age	0.0744** (0.0284)	0.0661** (0.0258)	0.0522** (0.0224)
Age sq.	-0.000824* (0.000438)	-0.000657 (0.000423)	-0.000599* (0.000332)
Males			
Age	-0.140* (0.0809)	-0.101 (0.0959)	-0.103* (0.0606)
Age sq.	0.00187* (0.00108)	0.00127 (0.00120)	0.00141* (0.000827)
Single female (omitted)			
Married female	-0.128 (0.112)	0.319 (0.205)	-0.263*** (0.0800)
Single male	3.630*** (1.127)	2.903* (1.671)	2.630*** (0.919)
Married male	3.994*** (1.052)	3.332* (1.867)	2.884*** (0.812)
Females			
Dependency ratio	-0.137*** (0.0401)	-0.273** (0.101)	
Dependency ratio x # Adult females	0.234** (0.0933)	0.382 (0.242)	
Males			
Dependency ratio	0.261** (0.112)	0.679*** (0.246)	
Dependency ratio x # Adult females	0.0126 (0.0758)	-0.166** (0.0715)	
Years schooling	0.136*** (0.00839)	0.0506*** (0.0125)	0.131*** (0.0102)
Enrolled in school	-0.569*** (0.165)	-0.201 (0.145)	-0.578*** (0.176)
Constant	-2.689*** (0.336)	-0.513 (0.365)	-1.929*** (0.283)
Rho			-0.288** (0.138)
Number of obs.	3,869		3,869

Note: standard errors robust to clustering at the city level in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The variable # of adult females counts the number of females aged 18+ in the household, excluding the survey respondent in the case of female respondents. City fixed effects included. Inference done with a t-distribution with # clusters -1 as degrees of freedom. All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table A4: Predicted log-earnings differential. Formal employment – actual. Formality status defined by job with social security

$\widehat{\ln y_f} - \overline{\ln y}$	Does not have social security	
	Wants social security	Does not want social security
	-0.947	1.013
	(.81)	(.604)

Note: predictions based on the models estimated in Table 10. Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).

Table A5: Predicted log-earnings differential. Formal employment – actual. Formality status defined by job with social security or formal self-employment

$\widehat{\ln y_f} - \overline{\ln y}$	Informal employment	
	Wants social security	Does not want social security
	-1.26	.462
	(.80)	(0.436)

Note: Predictions based on the models estimated in Table 11. Standard errors in parentheses. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . All estimates use sampling weights.

Source: author's calculations based on MOTRAL 2015 (INEGI 2015a) and ENOE 2<sup>nd</sup> Quarter 2015 (INEGI 2015b).